

W 1172



(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:  
24.10.2001 Bulletin 2001/43

(51) Int Cl.7: G06F 11/14

(21) Application number: 00117949.8

(22) Date of filing: 21.08.2000

(84) Designated Contracting States:  
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU  
MC NL PT SE  
Designated Extension States:  
AL LT LV MK RO SI

(30) Priority: 19.04.2000 JP 2000118012

(71) Applicant: Hitachi, Ltd.  
Chiyoda-ku, Tokyo 101-8010 (JP)

(72) Inventors:  
• Mimatsu, Yasuyuki, Hitachi, Ltd.,  
Intel.Prop.Group  
Chiyoda-ku, Tokyo 100-8220 (JP)

• Matsunami, Naoto, Hitachi, Ltd.,  
Intel.Prop.Group  
Chiyoda-ku, Tokyo 100-8220 (JP)

• Muraoka, Kenji, Hitachi, Ltd., Intel.Prop.Group  
Chiyoda-ku, Tokyo 100-8220 (JP)

• Oeda, Takashi, Hitachi, Ltd., Intel.Prop.Group  
Chiyoda-ku, Tokyo 100-8220 (JP)

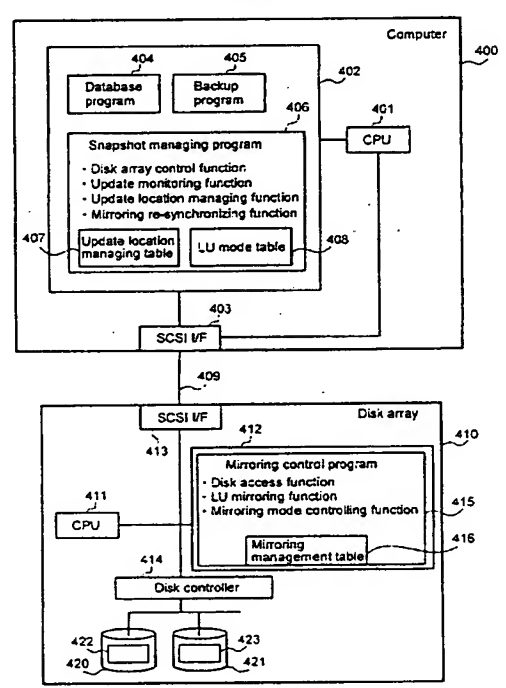
• Yagisawa, Ikuya, Hitachi, Ltd., Intel.Prop.Group  
Chiyoda-ku, Tokyo 100-8220 (JP)

(74) Representative: Strehl Schübel-Hopf & Partner  
Maximilianstrasse 54  
80538 München (DE)

(54) Computer system and snapshot data management method

(57) A method of managing a snapshot data of a computer system provided with a computer and a storage subsystem coupled with the computer, and the computer system to which the method is applied are disclosed. The storage subsystem is provided with the duplicated first and second storage units. In the state in which they are duplicated, when a request to update data is issued to the first storage unit, the storage subsystem writes the same updated data onto the second storage unit. When an acquisition of the snapshot is requested by the computer at any time, the storage subsystem suspends the writing of the writing data for the first storage unit onto the second storage unit thereafter. When the writing of the data onto the first storage unit occurs, the computer holds the location of writing the data as differential information. When making the contents of the first storage unit and the second storage unit the same, the data written onto the first storage unit after the acquisition of the snapshot is requested is written onto the second storage unit.

FIG.1



EP 1 148 416 A2

## Description

### Background of the invention:

[0001] The present invention relates to a data management method in a storage subsystem and to a computer system. The present invention relates, particularly, to the snapshot management method of an external storage unit necessary for the backup of the data of the external storage unit of the computer system of which high availability is required and to the computer system having the storage subsystem controlled by means of the above management method.

[0002] Generally, in the computer systems, the copies (backups) of data are made periodically on other recording medium such as magnetic tape in preparation for the loss of the data recorded in storage units such as hard disks caused by any of a failure of equipment, a bug in software, a wrong operation, etc. When the data recorded in the storage unit is lost, the original data can be recovered using the backed up data. When acquiring a backup, if the data is updated during a copying operation in order to acquire the backup, inconsistency of the copied data will be caused. Therefore, prevention of data update during the copying operation must be assured.

[0003] To avoid the update of the data being copied for backup, suspension of operation of all the programs except the backup program which accesses the data will do. However, in the system being required for high availability, the program cannot be suspended for a long time. Therefore, a mechanism which allows the data update during the backup operation and can keep the data at the time point of the backup start is necessary.

[0004] Generally, the fixed image of the data at a certain time point is called a snapshot and a mechanism which allows the update of the data while keeping the snapshot at the certain time point is called a snapshot management method. Taking the snapshot by means of the snapshot management method is called an acquisition of snapshot and the object data from which the snapshot is taken is called original data. The conventional snapshot managing method is realized by storing the pre-update data, duplication of the data by a computer or duplication of the data by an external storage unit. Conventional snapshot management method by means of the above is explained as follows.

#### (1) Method by means of storing the pre-update data

[0005] The method of snapshot management by means of storing the pre-update data is, for example disclosed by U.S. Patent No.5649152. In this disclosed technique, if the update of the original data occurs after the time point of the acquisition of the snapshot when acquiring a backup of certain data, the memory content before the update is stored in a different memory area. The snapshot is accessed as different data independent of the original data logically, but after the time point of

the acquisition of the snapshot data, the snapshot shares the memory area with the original data regarding the part in which the original data is not updated. Regarding the part in which the original data is updated, contents of the memory before the update stored in another area is referred to. In this case, the snapshot is exclusive to reading.

#### (2) Method by means of duplicating data by a computer

[0006] In the method by means of duplicating data by a computer, a program on the computer stores all the data in two areas duplicating (mirroring) the data, in the state in which the snapshot has not been acquired. When acquiring the snapshot, the program on the computer stops the duplicating operation separating the two memory areas to the independent regions, and provides one area as the original data and another as the snapshot. Such method of snapshot management is, disclosed by U.S. Patent No.5051887, for example.

#### (3) Method by means of duplicating data by an external storage unit

[0007] In the method by means of duplicating data by an external storage unit, as disclosed by U.S. Patent No. 5845295; for example, the management of the snapshot is executed in the procedure explained in Item (2) above inside the external storage unit instead of the program inside the computer. The entire function for managing the snapshot is in the snapshot management program inside the external storage unit.

### Summary of invention:

[0008] According to the above-mentioned prior art by means of the storage of pre-update data, accesses to the snapshot and those to the original data both refer to the same memory area concerning the parts in which the original data has not been updated after the acquisition of the snapshot. Therefore, it is feared that the accesses are concentrated at specific recording media, and the input/output performance of the disk is deteriorated. In this technology, both of the accesses to the original data and the snapshot are done via the snapshot management program. Therefore, during the backup, the load on the computer executing the snapshot management program is increased so that execution speed of other programs such as database programs is also feared to be slowed down. In the case of backing up a large amount of data, the performance is feared to be deteriorated for a long time till the end of the backup operation.

[0009] In the above-mentioned method by means of duplicating data by a computer, the snapshot management program writes data on two LU's in one data writing in the state in which the snapshot has not been acquired. That is, twice as many executions of the write operation

as those of the system without the snapshot management are required. Therefore, in the prior art, load on a CPU of the computer executing the data writing, the amount of data communication through communication paths connecting the computer and the external storage units, and the load on the disk controller of the external storage units are increased. For this reason, the execution speed of each application program will be slowed down compared with that in the system without the snapshot management. The deterioration of performance is conspicuous with the processing involving an update of a large amount of data such as the replication of database.

[0010] In the method by means of the prior art wherein the external storage unit duplicates the data, all the processings necessary for the snapshot management are installed in the external storage unit, so that the control programs of the external storage unit becomes complicated. The period required for development of the control program becomes longer corresponding to the complexity of the program, and thus the price of it rises, so that a problem is caused that the system becomes expensive.

[0011] The purpose of the present invention is to solve such problems of the prior arts and to provide the snapshot management method which can prevent the concentration of load on specific memory media of the external storage unit and the increase of load on the CPU of a computer in the state in which the snapshot of the external storage has been acquired, and alleviates the load on the CPU of the computer, an amount of communication through the communication paths, and the load on the external storage unit in the mode in which the snapshot of the external storage has not been acquired.

[0012] Another purpose of the present invention is to provide a computer system which is able to execute the snapshot management of the external storage using an inexpensive external storage unit having comparatively simple functions.

[0013] The present invention provides the method for managing the snapshot data in a computer system having a computer and a storage unit system connected to it and the computer system to which the method is applied. The storage unit system is equipped with the first and the second storage units which are to be duplicated. In the state in which they are duplicated, when a data update request is issued to the first storage unit, the storage unit system writes the same updated data on the second storage unit, too. When an acquisition of the snapshot is requested from the computer at any time, the storage unit system interrupts a writing data to be written on the first storage unit on the second storage unit thereafter. When a writing of data on the first storage unit occurs, the computer holds the writing location as differential information. When making the contents of the first storage unit and the second storage unit the same, the data written on the first storage unit after the acquisition of the snapshot was requested is written on

the second storage unit according to the differential information held in the computer. In one form of the present invention, the processing to make the contents of the first storage unit and those of the second storage unit the same is executed in the way that the computer reads the data from the first storage unit according to the differential information, and then instructs the storage unit system to write the data. In another form of the present invention, the processing to make the contents of the first storage unit and those of the second storage unit is executed in the way that the computer acquires the storage locations of the data to be made to be the same according to the differential information and instructs the storage unit to copy the data specifying the storage locations. In obedience to the copying instruction, the storage unit system reads out the data from the specified storage locations in the first storage unit and copies the data to the second storage unit.

[0014] The further features and the effects of the present invention will be understood by the following detailed explanations and drawings on the present invention.

Brief description of the drawings:

#### [0015]

Fig. 1 is a simplified block diagram showing the configuration of the computer system of a first preferred embodiment of the present invention.

Fig. 2 is a conceptual drawing showing the form of the mirroring management table.

Fig. 3 is a conceptual drawing showing the form of the update location management table.

Fig. 4 is a conceptual drawing explaining the contents of a parameter page used for Mode Select and Mode Sense.

Fig. 5 is a flowchart explaining the data writing operation of the snapshot management program.

Fig. 6 is a flowchart explaining the operation of the mirroring control program when it receives a write command from a computer.

Fig. 7 is a flowchart showing the operation of the mirroring re-synchronizing function of the snapshot management program.

FIG 8 is a simplified block diagram showing the configuration of the computer system of the second preferred embodiment of the present invention.

Fig. 9 is a conceptual drawing showing the form of the mirroring definition table.

Fig. 10 is a conceptual drawing showing the form of the LU mode table.

Fig. 11 is a conceptual drawing showing the contents of a parameter page to be returned to the Mode Sense command.

FIG 12 is a simplified block diagram showing the configuration of the computer system of the third preferred embodiment of the present invention.

Fig. 13 is a simplified block diagram showing the configuration of the fourth

Fig. 14 is a conceptual drawing showing the configuration of the mirroring definition table.

#### Description of the preferred embodiments:

[0016] Fig. 1 is a block diagram showing the configuration of a computer system in which the snapshot management of the first preferred embodiment of the present invention is executed. The first preferred embodiment of the present invention relates to the snapshot management in the case of executing the online backup of a storage area being accessed by a database program in parallel with the database processing. In the Fig. 1, the computer 400, CPU's 401 and 411, memories 410 and 412, SCSI interfaces 403 and 413, a database program 404, a backup program 405, a snapshot program 406, an update location management table 407, an LU mode table 408, an SCSI bus 409, a disk array 410, a disk controller 414, a mirroring control program 415, a mirroring management table 416, disk units 420 and 421, and logical units (LU's) 422 and 423 each of which is logically managed as a disk unit are illustrated.

[0017] The computer 400 and the disk array 410 are connected by the SCSI bus 409 via the interfaces 403 and 413. In the memory 402 inside the computer 400, the database program 404, the backup program 405, and the snapshot management program 406 which are executed by the CPU 401 controlling the computer 400 are stored. The disk units 420 and 421 which are controlled by the disk controller 414 are provided in the disk array 410. The mirroring control program 412 which is executed by the CPU 411 is stored in the memory 412. The storage areas inside each of the disk units 420 and 421 are accessed by the computer 400 as the logical units (LU's) 422 and 423 of SCSI.

[0018] The database program 404 has the function to access the LU 422 while being executed and to stop and restart the update of the data on the disk according to the user's instruction. The backup program 405 executes a reading out of data for a backup from the LU 423 according to the instruction by the user. The snapshot management program 406 is a middleware like the device driver and the volume manager. All the disk access requests issued by the database program 404 to the disk array 410 are issued to the disk array 410 via the snapshot management program 406. The snapshot management program 406 has the update location management table 407 and the LU mode table 408, and realizes the disk array control function, update monitoring function, update location management function, and mirroring re-synchronizing function.

[0019] The mirroring control program 415 of the disk array 410 has the LU mirroring function which writes the same data on another predetermined LU by duplicating an update of the one LU in addition to the disk access function which instructs the disk controller 414 to access

a disk in obedience to a request from the computer 400. Here, explanations will be made on the assumption that an update of the LU 422 is duplicated and applied to the LU 423. Those are functions that the conventional disk array has. Besides, the mirroring control program 415 has the mirroring mode control function to make each of the duplicated two LU's an accessible independent LU by invalidating the LU mirroring function.

[0020] The definition and the current mode of the mirroring mate LU of each LU to be duplicated are recorded in the mirroring management table 416. Fig. 2 is a drawing showing the form of the mirroring management table. The mirroring mate LU and its mode for each LU to be duplicated are recorded in the mirroring management table 416. The mode includes three types of them which are "Synchronized" in which contents of both LU's are the same, "Snapshot" in which each of the LU's can be accessed as an independent LU, and "Re-synchronized" in which the data which are not the same between the LU's are being copied to change the mode from the "Snapshot" to the "Synchronized". The LU mirroring function is applied to the LU in the mode of "Synchronized" or "Re-synchronized". By the setting of the two or more mirroring mate LU's, a multiple configuration of treble or more can be formed.

[0021] The snapshot management program 406 has, as described later, the disk array control function which issues the Mode Sense and Mode Select commands of SCSI (Small Computer System Interface), the update monitoring function which detects the update of the original data requested by the database program after the acquisition of the snapshot, the update location management function which records the location of the update, and the mirror re-synchronizing function which copies the updated part of the original data to the mirroring mate LU when the snapshot is deleted. The update monitoring function and the update location management function are applied to the LU in the mode of "Snapshot" or "Re-synchronized" and the mirror re-synchronizing function is applied only to the LU in the mode of "Re-synchronized".

[0022] The LU mode table 408 which records the current mode of each LU has the form of the mirroring management table 416 of the disk array shown in Fig. 2 from which the definition of the mirroring mate LU has been omitted. The LU mode table 408 is produced by the snapshot management program 406 when it issues the Mode Sense commands to all the accessible LU's and acquires the mode of each LU from the disk array 410 at the time when it is activated.

[0023] The update location of the original data after the acquisition of the snapshot is recorded in the update location management table 407. Fig. 3 is a drawing showing the form of the update location management table. The LU which stores the updated original data, the first logical block address (LBA) of the update location, and the last LBA of the update location are recorded in the update location management table. The up-

date location management table shown in Fig. 3 shows that the parts which are the LBA's 0 to 15 and the LBA's 256 to 511 of the LU 2, and the LBA's 8 to 31 of the LU 3 have been updated after the acquisition of the snapshot. The record in the update location management table 407 indicates the location where the contents of the two LU's which are defined as the duplicated pair are not the same.

[0024] Next, the control of the disk array from the computer in a computer system with above-mentioned configuration in the preferred embodiment will be explained as follows.

[0025] All the operation of the mirroring control program of the disk array for each LU such as validating/invalidating the LU mirroring function and permitting/rejecting access to the mirroring mate LU are decided according to the mode of the LU recorded in the mirroring management table. The mirroring control program 415 changes the mode recorded in the mirroring management table 416 to that specified by the Mode Select command issued by the snapshot management program 406. The mirroring control program 415 returns the current mode record of the specified LU when it receives the Mode Sense command which requires the contents of the mirroring management table 416. Fig. 4 shows the contents of a parameter page used for the Mode Select and Mode Sense commands. The parameter page comprises 2 bytes; a page code (00h) is specified at the top and the second byte indicates the mode of the specified LU. In the Mode Select command, one of 0, 1 and 2 is set as the mode corresponding to the "Synchronized", "Snapshot" and "Re-synchronized" respectively. When the Mode Sense command is executed, current mode is returned, however, if the LU specified by the Mode Sense command is not duplicated or defined as the mirroring mate LU, "4" is returned as "Undefined".

[0026] Next, the acquisition of the snapshot will be explained for the case of acquiring the snapshot of the LU 422 as an example.

[0027] In the case of acquiring the snapshot, the user stops an update of the database temporarily to assure the consistency of the data to be backed up. Then, the user instructs the snapshot management program to acquire the snapshot of the LU 422. The snapshot management program 406 refers to the LU mode table 408 and ascertains that the mode of the LU 422 is the "Synchronized". When doing that, if there is no record of LU 422 in the LU mode table or the mode is not the "Synchronized", the snapshot management program 406 discontinues the processing.

[0028] When the mode of the LU 422 is the "Synchronized", the snapshot management program 406 changes the mode recorded in the LU mode table 408 from "Synchronized" to "Snapshot" and validates the update monitoring function and the update location management function. Then, the snapshot management program 406 issues the Mode Select command to change the mode of the LU 422 recorded in the mirror manage-

ment table 416 inside the disk array 410 to "Snapshot". When the mirror control program 415 inside the disk array 410 receives the Mode Select command, it changes the mode of the LU 422 recorded in the mirror management table 416 to "Snapshot". Then, the user restarts the update of the database.

[0029] The operation of the mirroring control program 415 for each LU, that is, the control on reading/writing of data, is decided according to the modes of the LU as explained later. This change of the modes enables the LU 423 which is the destination of the mirroring mate LU to be accessed by a computer as an independent LU. Besides, since the update of the LU 422 is not reflected on the LU 423, the backup program 405 can use the LU 423 as the snapshot of the LU 422.

[0030] Next, the deletion of the snapshot will be explained for the case of deleting the snapshot of the LU 422 as an example.

[0031] In the case of deleting the snapshot, the user instructs the snapshot management program 406 to delete the snapshot of the LU 422. The snapshot management program 406 refers to the LU mode table 408 and ascertains that the mode of the LU 422 is the "Snapshot". The snapshot management program 406 discontinues the processing if there is no record of the LU 422 in the LU mode table 408 or the mode is not the "Snapshot". When the mode of the LU 422 is the "Snapshot", the snapshot management program 406 issues the Mode Select command to change the mode of the LU 422 recorded in the mirroring management table 416 inside the disk array 410 to the "Re-synchronized". This change of the mode makes the access from the computer destination of which is the LU 423 impossible. Thereafter, the LU mirroring function is validated and the update of the LU 422 is reflected on the LU 423.

[0032] The snapshot management program 406 changes the mode of the LU 422 recorded in the LU mode table 408 to the "Re-synchronized" and validates the mirroring re-synchronizing function. The details of the data copying processing performed by the mirroring re-synchronizing function will be explained later. When the data copying is completed, the snapshot management program 406 issues the Mode Select command which instructs the "Synchronized" and changes the mode recorded in the LU mode table 408 to the "Synchronized". The mirroring control program 415 of the disk array 410 changes the mode of the LU 422 recorded in the mirroring management table 416 to the "Synchronized" in obedience to the Mode Select command. Through the above-mentioned processings, the update monitoring function, the update location management function, and the mirroring re-synchronizing function of the snapshot management program for the LU 422 are invalidated and the deletion of the snapshot is completed.

[0033] The following is an explanation of the data reading out processing that is the reading out of the data in the LU 422 and the LU 423 of the disk array by the

database program 404 and the backup program 405 on the computer 400 respectively.

[0034] The input/output request from the program inside the computer 400 is issued via the snapshot management program 406. In the case where the input/output request is a reading request, the snapshot management program transfers the READ command of SCSI to the disk array. When the mirroring control program 415 of the disk array receives the READ command, it reads out data from the designated LU and returns the data and the status to the computer 400. When doing this, if the mode of the designated LU is the "Synchronized", the load is scattered by means of reading out from either the designated LU and mirroring mate LU because contents of the two LU's are the same. In the case where the designated LU is the mirroring mate LU and its mode is not the "Snapshot", the mirroring control program 415 rejects the access and returns an error status. The access to the mirroring mate LU, which is not in the mode of the "Snapshot" including the writing, may be rejected inside the snapshot management program 406.

[0035] Fig. 5 is a flowchart explaining the operation of data writing of the snapshot management program. This processing is executed when the database program 404 and the backup program 405 update the storage contents of the LU 422 and the LU 423 respectively.

[0036] When the snapshot management program 406 receives the writing request from the programs 404 and 405 (step 800), it ascertains whether the mode of the LU access destination LU is the "Snapshot" or not (step 801) by referring to the LU mode table 408. When the mode of the access destination LU is found to be the "Snapshot" as the result of judgment in the step 801, the snapshot management program 406 checks whether the update location of the access destination which is to be updated by the writing request has already been recorded in the update location management table 407 (step 802). If the update location of the accessing destination is not recorded on the update location management table 407, it records the location in the update location management table 407 and issues the WRITE command of SCSI. If it is already recorded, the program issues the WRITE command of SCSI leaving it as it is (steps 803 and 807).

[0037] If the mode of the LU is found to be not the "Snapshot" as the result of judgment in the step 801, the snapshot management program 406 further ascertains whether it is the "Re-synchronized" or not (step 804). Since the mode of the access destination LU is the "Synchronized" if the mode of the LU is not the "Re-synchronized", the program goes to step 807 leaving it as it is and issues the WRITE command. When the mode of the LU is found to be the "Re-synchronized" as the result of the judgment in step 804, the snapshot management program 406 checks further whether the update location to be accessed by the WRITE command is recorded in the update location management table 407 or not (step 805). If the information concerned is not recorded, the

program issues the WRITE command leaving it as it is. If the information on the update location is recorded, the program issues the WRITE command after deleting the concerning information on the update location from the update location management table 407 (steps 806 and 807).

[0038] When adding the information on the update location in the update location management table 407 in the above-mentioned processing, the information is added in the way that no multiple information including the same update location exists. When deleting the information of the update location from the update location management table 407, the update location management table 407 is amended in the way that no information other than that on the update location which is the object of the deletion is deleted. For example, the deletion of the information on the LBA=10 to 11 of the LU 3 from the table is assumed when the update location management table 407 has the contents shown in the Fig. 3. In this case, the LBA=8 to 31 of the LU 3 recorded in the update location management table 407 are deleted and an entry having the information on the LBA=8 to 9 and the LBA=12 to 31 is newly added.

[0039] Fig. 6 is a flowchart explaining the operation of the mirroring control program which has received the WRITE command from the computer.

[0040] When the mirroring control program 415 receives the WRITE command (step 900), it checks whether the access destination LU has been set as the mirroring mate LU or not (step 901) by referring to the mirroring management table 416. If the access destination is found to be the mirroring mate LU as the result of the judgment in the step 901, the mirroring control program 415 checks whether the mode is the "Snapshot" or not (step 905). If the mode of the LU is the "Snapshot", the program writes data on the LU designated in the step 904. If the mode of the LU is not the "Snapshot", the program returns the write failure status in the step 906 because the LU designated as the access destination cannot be accessed as an independent LU.

[0041] In the case where the access destination LU is not the mirroring mate LU, the mirroring control program 415 ascertains whether the mode of the LU is "Snapshot" or not (step 902). If the mode of the LU is not "Snapshot", the program goes to step 904 after writing data on the LU designated as the mirroring mate LU (step 903). The mirroring control program 415 writes the data on the LU designated in the step 904 and returns the status of the command execution to the computer 100 in the step 906.

[0042] In the processing of the snapshot management program 406 explained above, the reason for deleting the record in the update location management table 407 when the LU to be updated is in the "Re-synchronized" mode and the update location to be accessed by the write request is recorded in the update location management table 407 is to avoid the waste of data copying processing executed in the LU re-synchro-

nizing processing at the time of the snapshot deletion which will be explained later. In the mode of "Re-synchronized", the LU mirroring function of the mirroring control program 415 is valid. Therefore, when the data update is generated, the contents of the updated part of the LU is reflected on the mirroring mate LU and the both contents are the same. Thus, the data copying processing for the re-synchronization is not necessary. Since, the data copying for the re-synchronization is executed for the updated location recorded in the update location management table 407 to be explained later, the record of the location wherein both contents are the same is deleted from the update location management table 407.

[0043] Fig. 7 is a flowchart explaining the operation of the mirroring re-synchronizing function of the snapshot management program. The snapshot management program 406 copies the original data to the mirroring mate LU concerning the part wherein stored contents of the two LU's being defined as a pair for the duplication are not the same when deleting the snapshot. The mirroring re-synchronizing function of the snapshot management program 406 operates in parallel with other functions.

[0044] The snapshot management program 406 firstly checks whether registered information exists in the update location management table 407 or not when the processing of the mirroring re-synchronizing function is started. When the registered information is not in the update location management table 407, the snapshot management program 406 terminates the mirroring re-synchronizing processing (step 1001). When the registered information is in the update location management table 407, the snapshot management program 407 reads out one of the updated data held in the registered update location. This reading out is executed by an issue of the READ command of SCSI to the disk array (step 1003). When the data is returned from the disk array 410 as the response to the READ command issued at the step 1003, the snapshot management program 406 checks whether the information regarding the location of update being processed is deleted or not by referring again to the update location management table 407. Such a deletion occurs in the case where the database program updates the record at the same location immediately after the issue of the READ command (step 1004). When the information in all the update locations from where the data was read out are deleted as the result of the judgment, it means that contents of the two LU's were made to be the same owing to the update processing generated during the processing in step 1003 so that the data copying is not necessary. Therefore, in this case, the program returns to the processing in the step 1001 is returned continuing the processing of other updated data. On the other hand, the information of the update location concerning a part or all of the read out data is in the update location management table 407, these data of the update location are written back in the same location. The LU mirroring function is

valid in the disk array 410 during the re-synchronizing processing, and the same data is written also onto the mirroring mate LU by means of the writing back of the data on to the original location like this, thus the contents of the both are made to be the same (step 1005). The snapshot management program 406 deletes the information on the location of the data which has been written back from the update location management table 407. Then, the processing of step 1001 is resumed and above-mentioned processing is repeated till the completion of the processing for all the information on the update locations registered in the update location management table 407 (step 1006).

[0045] In the preferred embodiment of the present invention explained above, the locations of the data necessary to be copied are recorded only in the update location management table 407 inside the computer 400. Therefore, there is a possibility of losing the location information owing to a failure of the computer 400. In this case, the stored contents can be made to be the same without losing the updated contents by means of the copying of all of the original data to the mirroring mate LU.

[0046] In the first preferred embodiment of the present invention, since the LU to which the database program accesses and the mirroring mate LU used as the snapshot of the former LU are on different disks, the load of the disk accesses by each of the backup program and the database program can be scattered. As a result, the lowering of the disk access speed by the database program during the backup can be reduced.

[0047] Also in the first preferred embodiment of the present invention, the duplication processing of the data is executed by the mirroring control program in the disk array. Therefore, the computer needs to transmit the WRITE command to the disk array only once per data writing. Load on the CPU for the I/O processing and the amount of communications through the SCSI bus can be reduced, and the execution speed of the database program can be improved compared with the duplication executed by the computer.

[0048] Further, since the disk array of the first preferred embodiment of the present invention has neither the function to manage the update location during the snapshot nor the LU re-synchronizing function to be used in the deletion of the snapshot, control of the disk array is less complicated compared with that of executing all the data duplication and the snapshot management by the external storage unit and the manufacturing cost of the disk array can be reduced.

[0049] In the above-mentioned preferred embodiment of the present invention, the processing for the re-synchronization is performed by reading out and re-writing the inconsistent data, however, this processing can be simplified further. In concrete, the mirroring control program on the disk array is provided with a function to copy the data in the designated location between the duplicated LU's. The snapshot management program



on the computer issues the copying instruction to the disk array designating the location of the data instead of executing the processing of reading and writing the data. In this case, in the flowchart in Fig. 7, a step for issuing a predefined SCSI command for instructing the copying is provided instead of the processing in steps 1003 to 1005. When the mirroring control program receives this command, it copies the data in the designated location inside the disk array to the mirroring mate LU. Thus, the load on the CPU of the computer and the increase of the amount of communication through the SCSI bus owing to the data copying processing can be reduced.

[0050] Fig. 8 is a block diagram showing the configuration of the computer system of the second preferred embodiment of the present invention. The computer system of this preferred embodiment of the present invention is different from that shown in the Fig. 1 in the points that the snapshot management program 406 inside the computer 400 is newly provided with a mirroring definition table 1108 and that the mirroring control program 412 inside the disk array 410 is provided with a mirroring definition table 1116 instead of the mirroring management table. The other configurations are basically the same with those of the first preferred embodiment of the present invention except some functions, and the components are designated with the same reference numbers as those in Fig. 1. In this preferred embodiment of the present invention, the mode management of the LU which is executed by both of the computer and the external storage unit in the first preferred embodiment is executed by the computer only and the external storage unit is made to be stateless regarding to the snapshot management.

[0051] In Fig. 8, the mirroring control program 415 of the disk array 410 has the virtual LU forming function which provides the duplicated two LU's as one LU logically. The relations among the LU's handled by the virtual LU forming function is registered in the mirroring definition table 1116. Fig. 9 is a drawing showing the form of the mirroring definition table. The mirroring definition table 1116 shows the relations among the LU's forming a group comprising three LU's which are defined as a mirroring origin LU which has the original data, the mirroring mate LU with the duplicated contents thereof and a mirroring LU which handles these as one LU logically. All the three LU's are accessible from outside. The responsibility for keeping the contents of the mirroring origin LU and the mirroring mate LU the same is placed on the snapshot management program 406 which operates on the computer 400.

[0052] The snapshot management program 406 provides the LU being provided by the disk array 410 to other program as a virtual LU. The virtual LU is accessed by other programs in the same way as a real LU. When a program accesses the virtual LU, the real LU to be accessed is switched by the snapshot management program according to the mode of the virtual LU. Here, such

function is called the virtual LU forming function. The access destination and the status of the virtual LU is registered in the LU status table 408. Fig. 10 is a drawing showing the form of the LU status table used for this preferred embodiment of the present invention. The LU definition table 408 defines the READ destination LU to be accessed by the reading request designating the virtual LU and the WRITE destination LU to be accessed by the writing request designating the virtual LU corresponding to the virtual LU. As a status to be entered into the status column of the LU status table 408, there is the "Undefined" other than the statuses used for the LU status table of the first preferred embodiment of the present invention. The "Undefined" indicates the status that the virtual LU is not accessible.

[0053] For the update location management table 407, which has the similar form as the update location management table of the first preferred embodiment of the present invention, information indicating the virtual LU instead of the real LU is set as the information indicating the changed LU. In this preferred embodiment of the present invention, the object LU being monitored, managed, and accessed by the update monitoring function, the update location management function, and the mirroring re-synchronizing function of the snapshot management program respectively is also the virtual LU. Each of the functions is validated or invalidated according to the mode of the virtual LU similarly as the case of the first preferred embodiment of the present invention. All the functions are invalidated for the LU in the "Undefined" mode.

[0054] The snapshot management program 406 has the mirroring definition table 1108 having the same contents as that of the mirroring definition table 1116 inside the disk array. The snapshot management program 406 issues the Mode Sense commands to all the accessible LU's at the start up and creates the mirroring definition table 1108 based on the information returned from the disk array 410 as the response to the command. The disk array 410 returns a 4-byte parameter page having the structure shown in Fig. 11 to the Mode Sense command and informs of the content of the mirroring definition table 1116 regarding each of the LU's. A page code "0" is set for the first byte at the top of the parameter page. For the second byte of the parameter page, the classification of "0" is set if the LU designated by the Mode Sense is a mirroring LU and "1" is set otherwise. When the classification is "0", the values for identifying the mirroring origin LU and the mirroring mate LU are set for the third and the fourth byte of the parameter page respectively. The snapshot management program 406 creates an entry for the LU with the classification of "0", and registers the mirroring origin LU and the mirroring mate LU when creating the mirroring definition table 1108. The mirroring definition table 1108 is not necessarily created automatically using the Mode Sense, but may be created by the snapshot management program 406 when it is started by means of reading the contents



file defined by the user.

[0055] As the initial mode, the "Synchronized" is set as the mode of the virtual LU to be accessed by the database program 404, and the mirroring LU's of the mirroring definition table 1108 are set as both of the READ destination LU and the WRITE destination LU in the LU mode table 408. Besides, the "Undefined" is set as the mode of the virtual LU to be accessed by the backup program 405.

[0056] The following is the explanation of the disk access of the program on the computer in the second preferred embodiment of the present invention.

[0057] The disk access from the program on the computer is executed designating the virtual LU provided by the snapshot management program 406. When the snapshot management program 406 receives the access request, it refers to the LU mode table 408, ascertains whether the mode of the designated virtual LU not the "Undefined", and rejects the access if the mode is "Undefined". Then, the snapshot management program 406 checks the designated READ destination LU and the WRITE destination LU of the designated virtual LU by referring to the LU mode table 408 and decides the real LU to be the access destination of the command to be issued. For example, in the mode shown in Fig. 10, the snapshot management program 406 issues a SCSI command targeting the LU 0 together with the reading and the writing requests regarding the access request designating the virtual LU 0. On the other hand, when receiving the access request designating the virtual LU 4, the snapshot management program 406 issues the command targeting the LU 7 in the case of the reading request and LU 6 in the case of the writing request. The snapshot management program 406 can switch the LU of the access destination LU with the virtual LU designated by each of the programs as the access destination fixed by rewriting the LU mode table.

[0058] When the mirroring control program 415 of the disk array 410 receives the access command, it executes the access to the corresponding LU referring to the mirroring definition table 1116. For example, in the case where the mirroring definition table is defined as shown in Fig. 9 when the mirroring control program 415 receives the access command targeting the LU 0, it judges the LU 0 the mirroring LU comprising the two LU's by referring to the mirroring definition table. Then, the mirroring control program 415 writes the same contents on the LU 1 and LU 2 using the LU mirroring function in the case of the update, and reads out data from either the LU 1 or LU 2 in the case of the reading out. On the other hand, the mirroring control program 415 accesses either the LU 1 or LU 2 corresponding to the access command designating the LU 1 or LU 2.

[0059] The following is the explanation of the processing for acquiring or deleting the snapshot in this preferred embodiment of the present invention. In the first preferred embodiment of the present invention, the snapshot management program managed the snapshot

by issuing the Mode Select commands to the disk array when acquiring or deleting the snapshot. On the contrary, the snapshot is managed through a change of the access destination LU of the virtual LU recorded in the LU mode table 408 in this preferred embodiment of the present invention.

[0060] The mirroring LU of the mirroring definition table is set up for both of the READ destination LU and the WRITE destination LU in the LU mode table 408 before the acquisition of the snapshot. At the time of the acquisition of the snapshot, the snapshot management program 406 rewrites those with the origin mirroring LU recorded on the mirroring definition table. The snapshot management program 406 sets the READ destination LU and the WRITE destination LU corresponding to the virtual LU of the snapshot to which the backup program 405 accesses to the mirroring mate LU and sets the mode to "Snapshot". For example, consider the acquisition of the snapshot of the virtual LU 0 in the modes shown in the Fig. 9 and Fig. 10. In this case, since the snapshot management program 406 can know that the mirroring origin LU of the LU 0, which is the current access destination, is LU 1 by referring to the mirroring definition table, it rewrites the READ destination LU and the WRITE destination LU of the virtual LU 0 with the LU 1 and sets its mode to "Snapshot". The snapshot management program 406 rewrites the READ destination LU and the WRITE destination LU of the virtual LU 1 to be accessed as the snapshot of the virtual LU 0 with the LU 2 which is the mirroring mate LU of the LU 0 and sets its mode to "Snapshot". With the above, an access to the virtual LU 1 is enabled. Because the update of the virtual LUs 0 and 1 cannot be duplicated each other, each of the virtual LU's becomes able to be utilized for the access to the original data and the snapshot.

[0061] When deleting the snapshot, the snapshot management program 406 changes the mode of the virtual LU of the snapshot set in the LU mode table 408 to "Undefined" and that of the virtual LU of the original data to "Re-synchronized". The WRITE destination LU of the virtual LU of the original data is changed to the mirroring LU which is registered in the mirroring definition table. For example, when deleting the virtual LU 3 which is the snapshot of the virtual LU 2 in the mode shown in Fig. 9 and Fig. 10, the snapshot management program 406 sets the mode of the virtual LU 3 to "Undefined". Since it can be known that the mirroring LU of the LU 4 which is the current WRITE destination LU of the virtual LU 2 is LU 3 through the reference to the mirroring definition table, the snapshot management program 406 rewrites the WRITE destination LU from the LU 4 to the LU 3 and sets its mode to "Re-synchronized". Through the above, the access from the program to the virtual LU 3 which is the snapshot becomes impossible. When performing the data copying processing for making the contents of the mirroring origin LU and the mirroring mate LU the same, each access destination is the mirroring origin LU for the reading out from the virtual LU of the original data

or the mirroring LU for the writing back of the original data onto the virtual LU. After the completion of the data copying processing, the snapshot management program 406 changes the READ destination LU of the virtual LU to be accessed as the original data to the original mirroring LU. To complete the deletion of the snapshot of the virtual LU 4 in the "Re-synchronized" mode, the snapshot management program 406 changes the READ destination LU to LU 6 knowing that the mirroring LU of the LU 7 which is the READ destination LU is the LU 6 from the mirroring definition table, and set its mode to "Synchronized". Thus, the deletion of the snapshot is completed.

[0062] According to the second preferred embodiment of the present invention, the recovery processing against a failure can be simplified by a transformation of the external storage unit to stateless mode. For example, if the computer fails at the moment the mode of the LU is changed in the first preferred embodiment of the present invention, the mode of each LU recorded in the LU mode table of the computer must agree with that recorded in the mirroring management table of the disk array in the recovery processing. However, in the second embodiment of the present invention, since the disk array is stateless, the processing can be preferred taking only the mode of the computer in consideration.

[0063] Fig. 12 is a block diagram showing the configuration of the computer system of the third preferred embodiment of the present invention. In this preferred embodiment of the present invention, each of the database program and the backup program runs on a different computer. In Fig. 12, a database server 1510 is a computer on which a database program 1515 which performs database management runs and a backup server 1520 is a computer on which a backup program 1526 runs. The database server 1510 and the backup server 1520 are connected mutually via a LAN (Local Area Network) 1500. The disk array 1550 which stores the database is shared by the database server 1510 and the backup server 1520. Among each of the servers and the disk array 1550 are connected via the Storage Area Network (SAN) comprising a fibre channel switch 1530 and fibre channel cables 1531, 1532, and 1533. A LAN 1500 is connected with a management console 1540 which is a computer on which a backup instruction program 1542 for instructing the start and stop of the backup runs via a LAN interface 1541.

[0064] The database server 1510 has a CPU 1511, a memory 1512, an LAN interface 1513 for connecting with the LAN 1500 and a fibre channel interface 1514 for connecting with the SAN. In the memory 1512, a communication program 1511 and a snapshot management program 1517 are installed together with the database program 1515. The database program 1515 and the snapshot management program 1517 execute processings similar to those of the programs of the preferred embodiment of the present invention explained above. The communication program 1516 receives the

instruction concerning a backup processing and calls the database program or the snapshot management program according to the contents thereof by communicating with the management console 1540 via the LAN 1500.

[0065] The backup server 1520 has a CPU 1521, a memory 1522, a LAN interface 1523, and a fibre channel interface 1524. In the memory 1522, a communication program 1525 which calls a backup program 1526 based on the instruction from the management console 1540 is installed together with the backup program 1526. The backup program 1526 accesses the disk array 1550 designating a real LU provided by the disk array 1550.

[0066] The functions of the disk array are the same as those of the second preferred embodiment of the present invention excepting that the communication is replaced with a SCSI protocol defined on the fibre channel.

[0067] The acquisition processing of the backup utilizing the snapshot management of this preferred embodiment of the present invention is performed in the following procedure.

[0068] When the backup instructing program 1542 of the management console 1540 receives the instruction to acquire the backup from the user, it instructs the communication program 1516 of the database server 1510 to acquire the snapshot. When the communication program 1516 receives the instruction, it instructs the database program 1515 to suspend the update operation temporarily and instructs the snapshot management program 1517 to acquire the snapshot. The acquisition of the snapshot is done by the snapshot management program 1517 in the same way as that of the second preferred embodiment of the present invention. After acquiring the snapshot, the communication program 1516 instructs the database program to restart the update of the database and reports the completion of the acquisition of the snapshot to the backup instructing program 1542.

[0069] When receiving the report of the completion of the snapshot acquisition, the backup instructing program 1542 instructs the communication program 1525 of the backup server 1520 to start the backup. When the communication program 1525 receives the instruction, it instructs the backup program 1526 to start the backup and makes it execute the backup. After completing the backup, the communication program 1526 reports the completion of the backup to the backup instructing program 1542.

[0070] When receiving the report of the backup completion, the backup instructing program 1542 instructs the communication program 1516 to delete the snapshot. When receiving the instruction, the communication program 1516 instructs the snapshot management program 1517 to delete the snapshot. After deleting the snapshot, the communication program 1516 reports the completion of the snapshot deletion to the backup in-

structing program 1542. Through the above processings, the online backup of the database is completed.

[0071] With the third preferred embodiment of the present invention, the load on the CPU of the database server during the backup can be reduced in addition to the effect obtained by the second preferred embodiment of the present invention. In the second preferred embodiment of the present invention, both the backup program and the database program run on the same computer. Therefore, owing to the increase of the load on the CPU caused by the data transferring processing during the backup, the execution speed of the database program is feared to be slowed down. Whereas, in the third preferred embodiment of the present invention, the backup program runs on a computer other than that for the database server so that the slowing down of the execution speed of the database program during the backup is prevented because the load of the data transfer processing does not affect the load of the CPU of the database server. With the third preferred embodiment of the present invention, quick backup can be realized by virtue of the high-speed data transfer because of the connection between the backup server and the disk array with the fibre channel.

[0072] Fig. 13 is a block diagram showing the configuration of a computer system of the fourth preferred embodiment of the present invention. This preferred embodiment of the present invention makes the substantial change of the conventional snapshot management program unnecessary when shifting the method of snapshot management from that by means of storing the pre-update data explained as a prior art to that by means of the above explained first or second preferred embodiment of the present invention. For this purpose, in this preferred embodiment of the present invention, the original data is stored in a mirroring LU in which the stored contents are duplicated on the two different memory media, and the original data is updated after the pre-update data is stored in another area when the update is requested after the time point when the snapshot has been acquired. The snapshot is represented as a virtual LU by the snapshot management program, but the part not updated is accessed referring to the original data and the updated part is accessed referring to the pre-update data stored in another area. The basic operation of the snapshot management program of this preferred embodiment in the present invention is similar to that used in the method for managing the snapshot of the prior art by means of storing the pre-update data, however, the function of distributing the access destinations, which will be explained later, is added for cooperation with the disk array supporting function.

[0073] The configuration of the hardware of the computer system shown in Fig. 13 and the functions of the database program and backup program are similar to those of the first and second preferred embodiment of the present invention explained above. The mirroring control program 415 which operates inside the disk ar-

ray 410 has the mirroring definition table 1116 which defines the relations among LU's similarly to the second preferred embodiment of the present invention. Fig. 14 is a drawing showing the form of the mirroring definition table of this preferred embodiment of the present invention. The mirroring definition table 1116 defines only the relation of correspondence between the mirroring LU and the mirroring mate LU different from the mirroring definition table of the second preferred embodiment of the present invention. The mirroring LU is the duplicated two LU's represented as a virtual LU similarly to that of the second preferred embodiment of the present invention and the mirroring mate LU is one of the two LU's comprising the mirroring LU. In the mirroring definition table 1116 shown in Fig. 14, the LU 0 is duplicated to the LU 1, for example. When the mirroring control program 415 receives the access command to the LU 0, it reads the data from either of the LU 0 or the LU 1 when the command is the reading request or writes the same contents to both of the LU 0 and the LU 1 when the command is the writing request. The access to the LU 1 which is the mirroring mate LU is limited only to the LU 1 and the LU 0 cannot be accessed. The LU which is registered as a mirroring mate LU in the mirroring definition table 1116 is a read only LU and the writing request to the mirroring mate LU is rejected by the mirroring control program 415. Through the above, the contents of the mirroring origin LU and the mirroring mate LU are always made to be the same. The contents of the data read from the mirroring LU and those read from the mirroring mate LU are the same, but when the mirroring mate LU is accessed, the access destination LU is limited to one LU. Application programs such as the database program 404 on the computer always access the mirroring LU. The mirroring mate LU is provided for the access destination allocation function of the snapshot management program 406 to be explained later.

[0074] The snapshot management program 406 is provided with the function of allocating a part of the access commands issued to the mirroring LU to the mirroring mate LU in order to refer to the original data. The mirroring mate LU corresponding to the mirroring LU of the original data is obtained through the reference to the mirroring definition table 1108. The mirroring definition table 1108 has the same contents as those of the mirroring definition table 1116 inside the disk array 410. The mirroring definition table 1108 is created from data acquired by the snapshot management program 406 from the disk array 410 with the Mode Sense, or by a manual setting of the users.

[0075] Following is the explanation of the access destination allocation function of the snapshot management program 406.

[0076] Access to the original data are classified into two kinds. One is the access from the database program 404 to the original data. The other is the access from the backup program 405 to the snapshot and it is the access in case the data of the access destination has

not been updated after the time point of the acquisition of the snapshot. In this case, the snapshot shares the memory area with the original data, so that the accessing destination is the original data. According to the prior art, both of those kinds of access are done to the mirroring LU being provided for the access to the original data, but in this preferred embodiment of the present invention, the accessing destination of the latter is changed to the mirroring mate LU by the snapshot management program 406. Thus, the access destinations of all the read requests to the original data generated from the backup program 405 are limited to the mirroring mate LU. Therefore, the access is loaded only on the magnetic disks to which the mirroring mate LU is allocated.

[0077] On the other hand, the access destination of the read requests generated by the database program 404 is the mirroring LU, and either of the duplicated two LU's can be used. Therefore, by using the disk bearing lighter load as the access destination for the reading request, the disk access speed of the database program is not lowered even when the load of the reading out, which is generated from the backup program, of the original data is heavy. For example, in the mirroring definition table 1108 shown in Fig. 14, the LU 0 and LU 1 are assumed to be allocated to the disk units 420 and 421 shown in Fig. 13 respectively. In this case, when the snapshot of the LU 0 is acquired, all the access requests from the backup program to the original data are issued to the LU 1 and only the disk unit 421 is used. Since the access requests issued by the database program 404 are pointed at the LU 0, both of the disk unit 420 and 421 can be used. Therefore, if the mirroring control program 415 selects the disk unit 420 bearing lighter load as the access destination, the disk reading out performance of the database program will not be deteriorated even when the disk access load caused by the backup is heavy.

[0078] The fourth preferred embodiment of the present invention can be realized by only adding the function to allocate the access destinations to the snapshot management program of the prior art and the period required for the development of the snapshot management program can be shortened.

[0079] According to each of the preferred embodiments of the present invention explained above, since the storage area used for the access to the original data and that used for the access to the snapshot are on the different memory media, the concentration of the load on a specific medium during the backup can be avoided. Besides, the backup processing can be executed on a computer other than that for online processings, and thus the increase of the load on the CPU during the backup processing can be lessened. Further, since the duplication of the data writing is executed by the external storage unit, the additional load on the CPU or the amount of the data communication caused by the duplication process of the data writing are not generated. The

external storage unit has neither the function for managing the update locations during the acquisition of the snapshot nor the function for assuring the consistency of the duplicated storage contents when the snapshot is deleted. Therefore, the cost of the external storage unit can be reduced and inexpensive computer system can be constructed.

## 10 Claims

1. A method of managing snapshot information in a computer system comprising a computer and a storage subsystem, which is coupled with said computer, having two or more storage units, said method comprising the steps of:

holding, in said storage subsystem, duplicated data in a first and a second storage units being correlated with each other in advance, and writing data on both of said first and said second storage units, in obedience to data writing requests from said computer;  
 setting the mode of said storage subsystem to the mode for acquiring said snapshot;  
 executing, by said storage subsystem, write operation to write data into said first storage unit in accordance with a write request issued by said computer while said storage subsystem is in the mode for acquiring said snapshot;  
 holding, by said computer, the storage location of data to be updated according to the write request issued after said setting step as differential information; and  
 setting the mode of said storage subsystem to the mode for getting re-synchronized, and making contents of said first storage unit and contents of said second storage unit the same based on said differential information.

2. The method according to claim 1 further comprising a step of writing, by said storage subsystem, the write data according to said writing request into both of said first and said second storage units when the write request designating said first storage unit is received while its mode is in said mode for getting re-synchronized.
3. The method according to claim 1, wherein said making step comprising the steps of:

acquiring, by said computer, the location of the data being updated according to the write request issued after the acquisition request of said snapshot based on said differential information and reading out the data stored in said location from said storage system;  
 requesting, by said computer, said storage sub-

system to write the data read out based on said differential information; and  
storing, in response to said request, the data read out based on said differential information in said second storage unit by said storage subsystem.

4. The method according to claim 1, wherein said making step comprising the steps of:

acquiring, based on said differential information, the location of data which has been updated by the write request issued after the acquisition request of snapshot;  
requesting said storage subsystem to copy the data with designating the acquired location of said data; and  
copying, by said storage subsystem, the data in the location designated by said request from said first storage unit to said second storage unit.

5. The method according to claim 1, wherein said making step includes a step of setting the mode in which the access designating said second storage unit is prohibited in response to said instruction.

6. The method according to claim 1, further comprising a step of setting a mode of said second storage unit to prohibit writing write data into said second storage unit in response to the request for acquiring said snapshot.

7. The method according to claim 1, further comprising a step of copying the data being held in said second storage unit to said third storage unit, and wherein said making step is performed after the completion of said copying step.

8. The method according to claim 7, wherein said copying step is performed by another computer with which said storage subsystem is coupled.

9. A computer system comprising:

a computer; and  
a storage subsystem being coupled with said computer, said storage subsystem including a first storage unit, a second storage unit, and a control unit controlling said first and said second storage units, where said control unit writes the write data accompanying said write request into both of said first and second storage units in accordance with the request, from said computer, to write data into said first storage unit when said control unit is set in the synchronized mode by said computer, and prohibits writing data on said second storage unit when it is set

in the snapshot mode by said computer; and said computer including means which selectively sets a operation mode including said synchronized mode and said snapshot mode to said storage subsystem, and means which holds the differential information indicating data writing location according to said writing request when the data writing request is issued to said storage subsystem while said storage subsystem is in a snapshot mode.

10. The computer system according to claim 9, wherein said operating mode includes the re-synchronized mode for making contents of said first storage unit and said second storage unit the same, and said control unit prohibits the reading out processing from said second storage unit when it is set in the re-synchronized mode by said computer.

11. A computer system according to claim 10, wherein said computer has means for acquiring, based on said differential information, a storage location in said first storage unit to which data was written when said storage subsystem was in the snapshot mode, reading out the data from the acquired storage location in said first storage unit, and issuing write request to write the read out data into said storage location.

12. A computer system according to claim 10, wherein said computer acquires, based on said differential information, the storage location in said first storage unit to which data was written when said storage unit was in the said snapshot mode, and instructs said storage subsystem to copy data held in said storage location, and wherein said control unit copies the data held in said storage location in said first storage unit to said second storage unit in response to said copying instruction.

13. A computer system according to claim 10, wherein said control unit rejects read request designating said second storage unit when it is set in said re-synchronized mode.

14. A computer system according to claim 9, wherein said control unit rejects write request designating said second storage unit when it is set in said snapshot mode.

15. A computer system according to claim 9, further having a storage device and wherein said computer copies the contents of said second storage unit to said storage device when said storage subsystem is in said snapshot mode.

16. A computer system according to claim 9, further comprising:

a storage device; and  
 a backup computer being communicably coupled with said computer;  
 wherein said backup computer receives a notification that said storage subsystem is in said snapshot mode from said computer and copies the data held in said second storage unit to said third storage unit in response to said notification.

10

17. A computer system, comprising:

a storage subsystem including a first storage unit, a second storage unit, and a control unit controlling said first and said second storage units; and  
 a computer having:

15

means for selectively setting operating mode, which includes a synchronized mode and a snapshot mode, to said storage subsystem,  
 means for receiving write request to said storage subsystem issued by an application program running on the computer,  
 means for issuing a write request with designating a first identifier discriminating a pair of said first and said second storage units when said storage subsystem is set in said synchronized mode by said setting means for said storage subsystem, and for issuing write request designating a second identifier discriminating said first storage unit when said storage subsystem is set in said snapshot, and  
 means for acquiring and holding differential information indicating location of said first storage unit designated by write request designating said first identifier.

20

25

30

35

40

18. A computer system according to claim 17, wherein said control unit writes data accompanying said writing request on both of said first and said second storage units in response to the write request designating said first identifier.

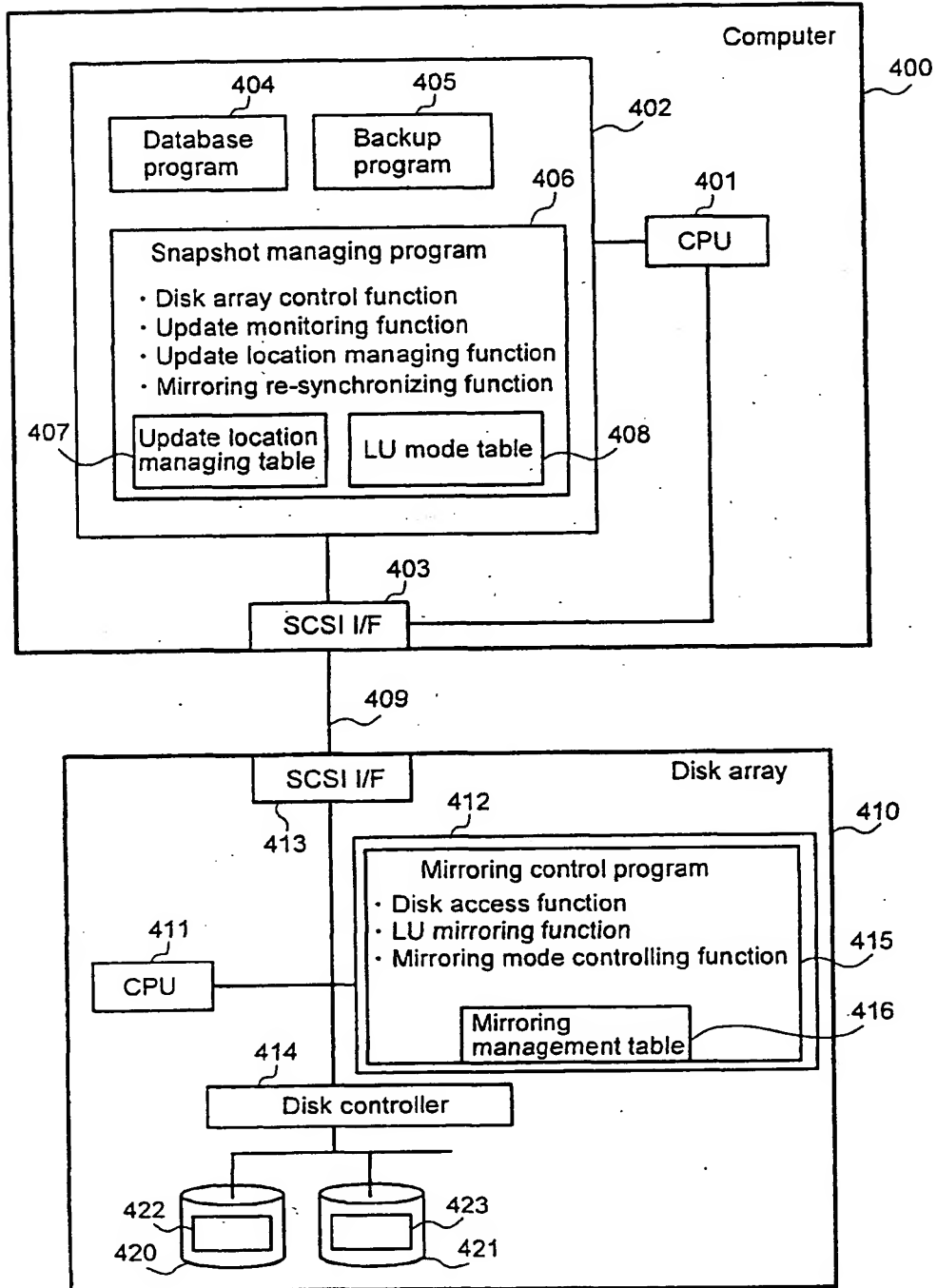
45

19. A computer system according to claim 18, wherein said operation mode includes the re-synchronized mode which makes the contents of said first storage unit and said second storage unit the same, and said issuing means issues the write request designating said first identifier in obedience to write request from a program which runs on said computer when said storage subsystem is in said re-synchronized mode.

50

55

FIG.1





**FIG.2**

415

LU	Mirroring destination LU	Mode
1	4	Synchronized
2	5	In snapshot
3	6	Re-synchronizing

**FIG.3**

407

LU	Update part starting LBA	Update part stopping LBA
2	0	15
2	256	511
3	8	31

**FIG.4**

Page code 00h
Mode

FIG.5

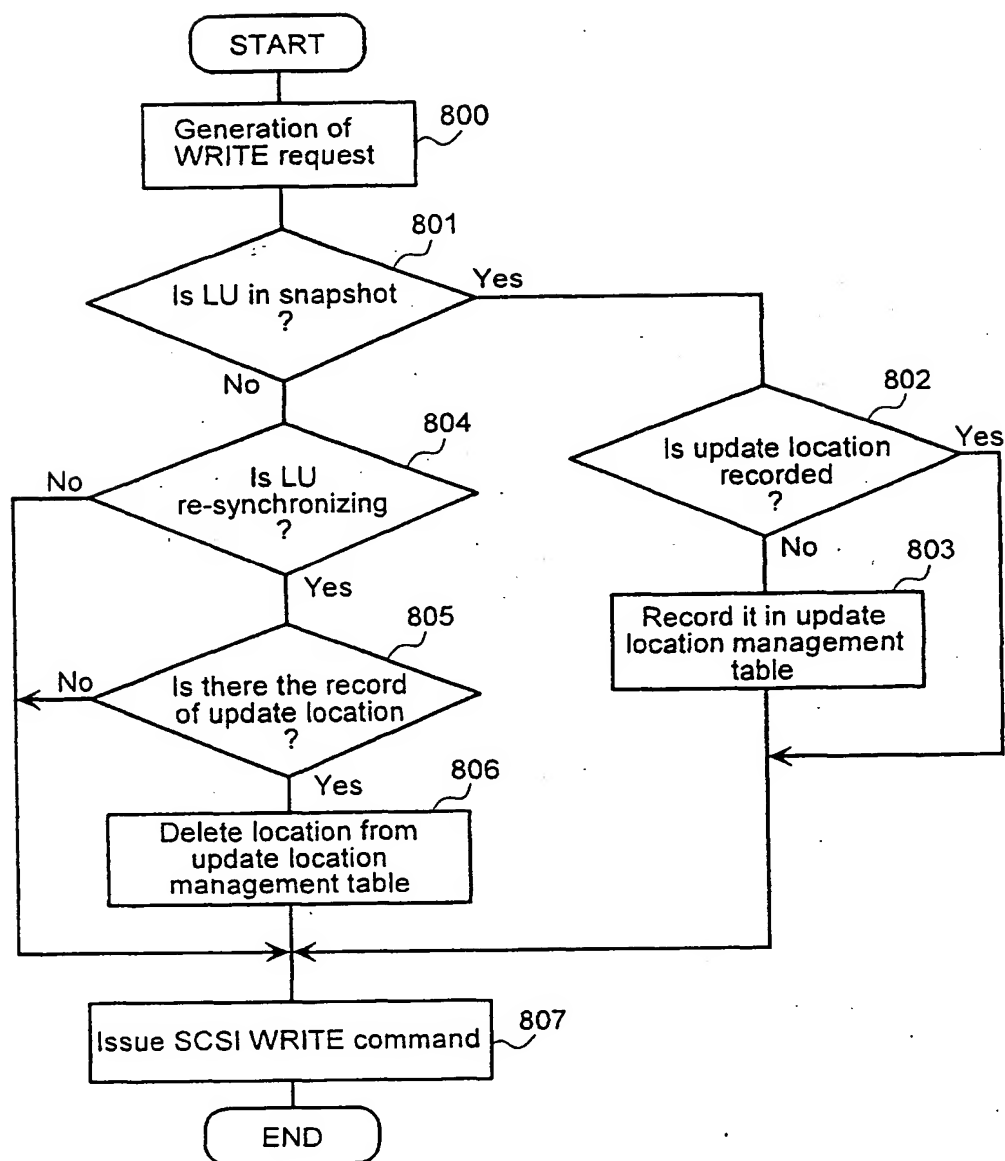


FIG.6

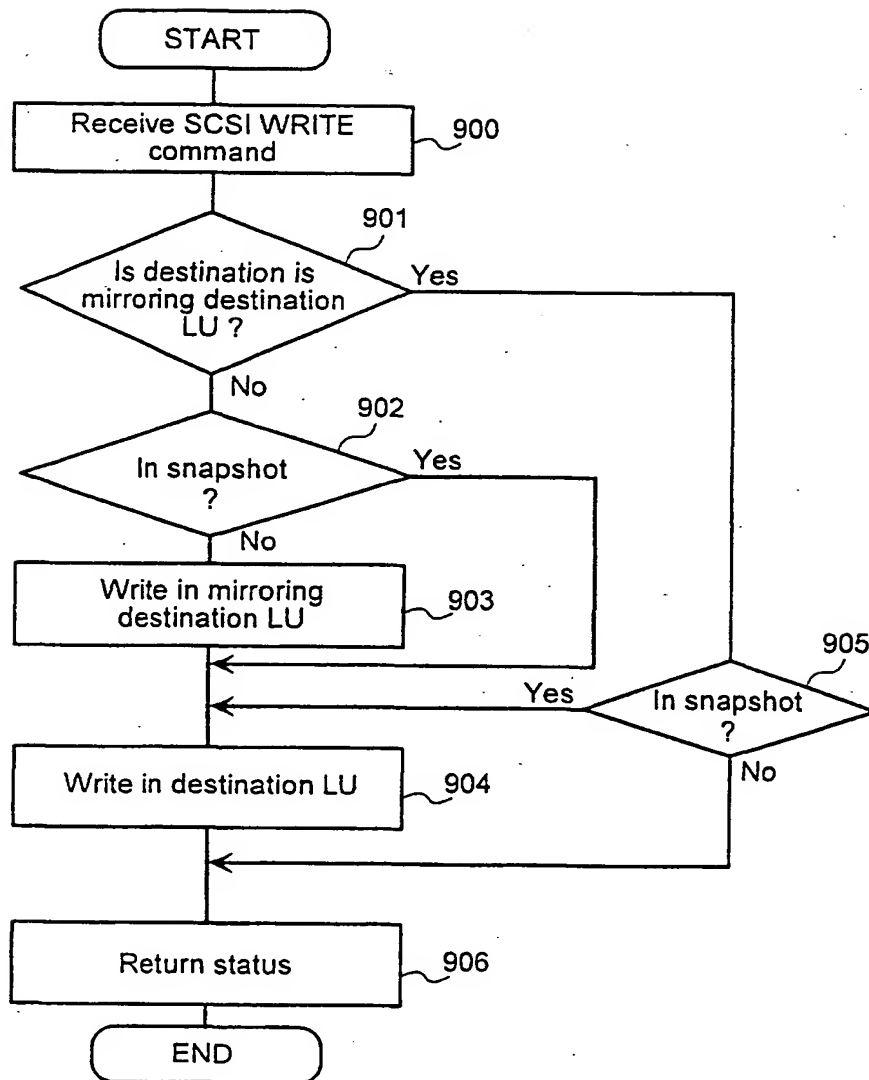


FIG.7

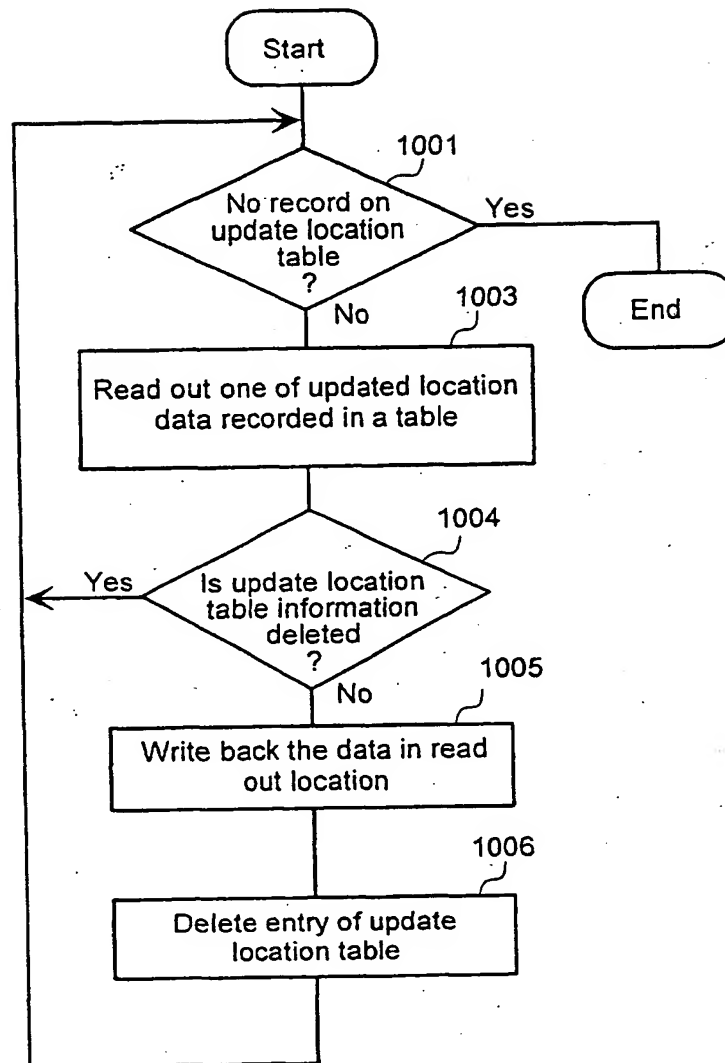
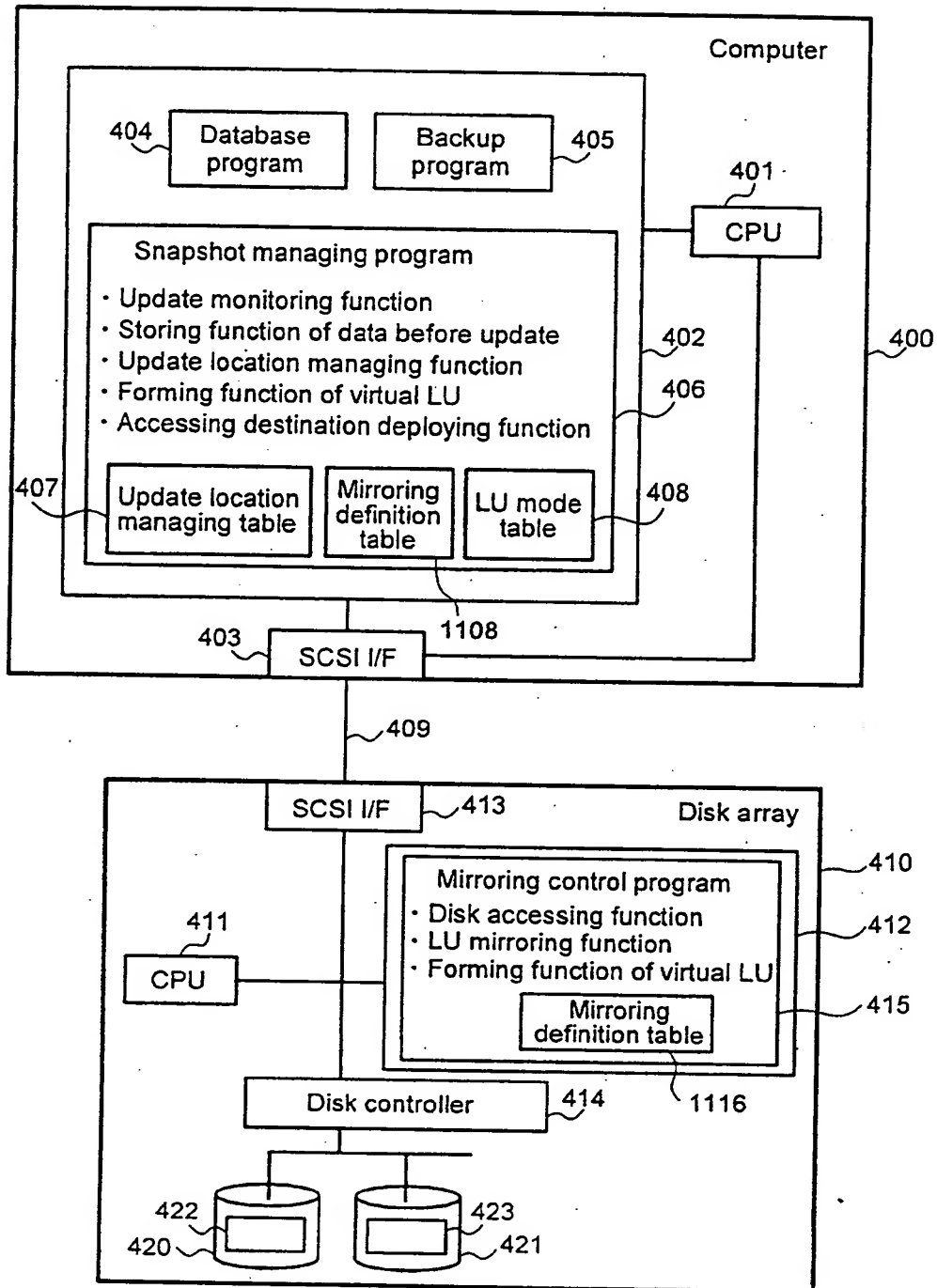


FIG.8



**FIG.9**

1116

Mirror LU	Mirroring origin LU	Mirroring destination LU
0	1	2
3	4	5
6	7	8

**FIG.10**

408

Virtual LU	READ destination LU	WRITE destination LU	Mode
0	0	0	Synchronization
1	—	—	Undefined
2	4	4	In snapshot
3	5	5	In snapshot
4	7	6	Re-synchronizing
5	—	—	Undefined

**FIG.11**

Page code 00h
Classification
Mirroring origin LU
Mirroring destination LU

FIG.12

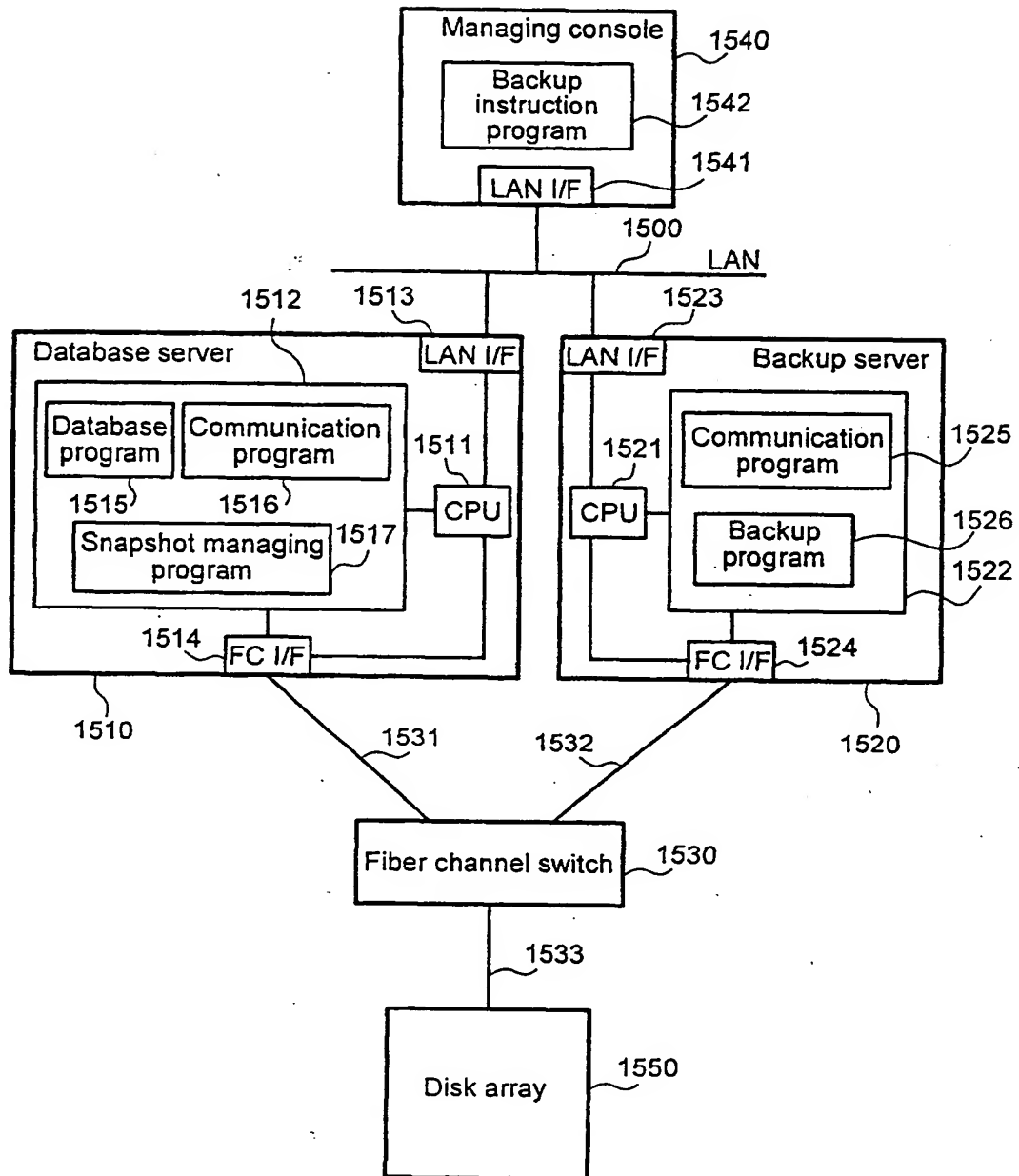
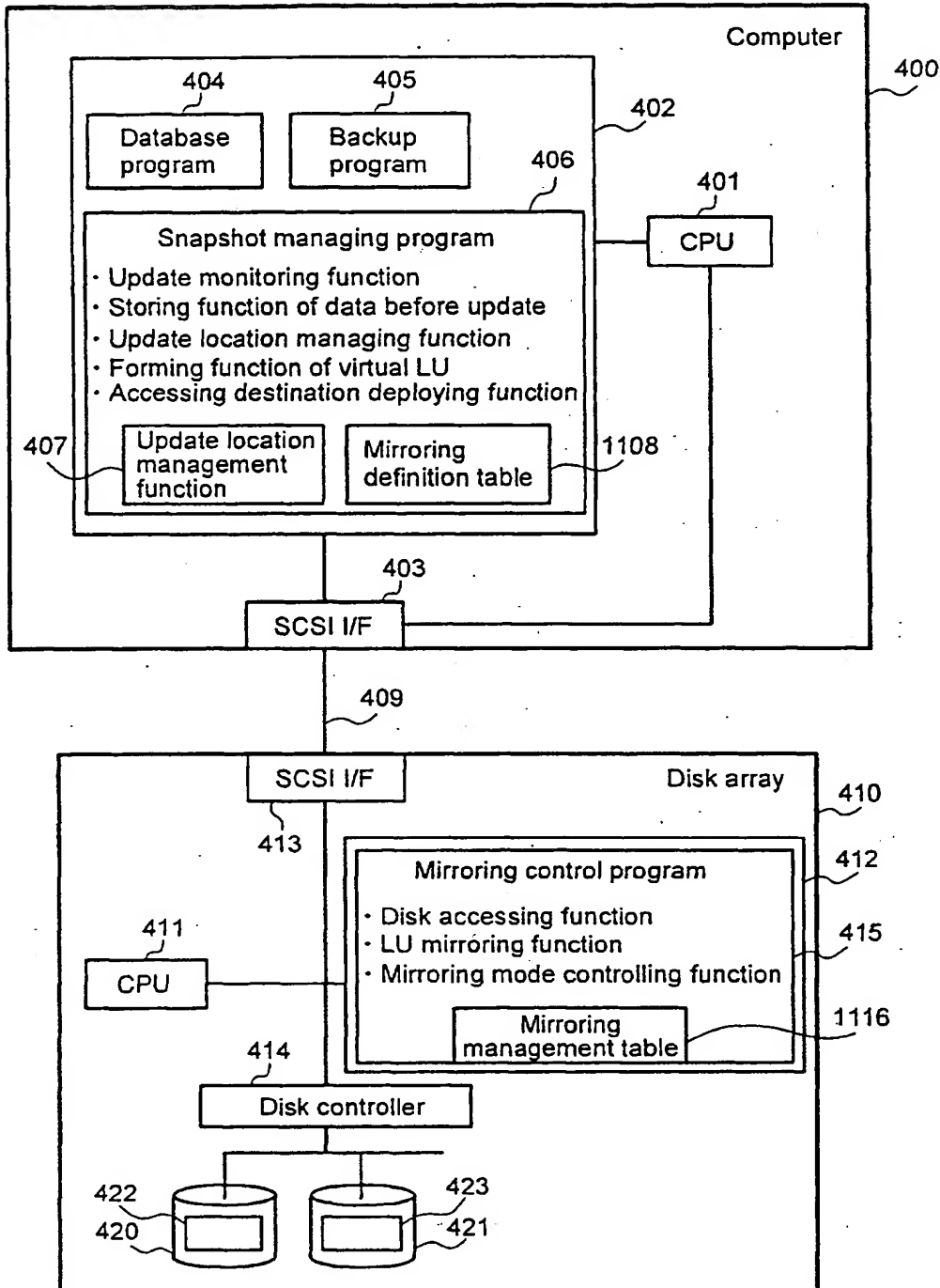




FIG.13



**FIG.14**

Mirroring LU	Mirroring destination
0	1
2	3
4	5